



Using field data in phonetics classes

Seunghun J. Lee^{1, 2}

¹International Christian University, Japan, ²University of Venda, South Africa,
¹seunghun@icu.ac.jp

Abstract

All phonetics classes teach acoustic phonetics that requires that learners develop an understanding of various concepts such as voice onset time and formants, build skills to use Praat for analyzing data, and apply phonetic knowledge in understanding sound systems of understudied languages.

This paper showcases three understudied languages for teaching these concepts. Voice onset time is introduced with Dzongkha data that has a four-way laryngeal contrast. Contrastive duration in Miyakoan Ryukyuan helps learners to understand the acoustics of singleton versus geminates. Vowels in Xitsonga are used for demonstrating the relationship between formants and vowel qualities in a five-vowel system.

Keywords: Teaching phonetics, field data, Dzongkha, Miyakoan Ryukyuan, Xitsonga

1. Introduction

Typical undergraduate phonetics classes are structured so that students learn phonetic concepts, practice how to use software such as Praat, and apply the knowledge in understanding phonetic theories. Teaching phonetics can be challenging because courses need to cover the physiology of articulation, acoustic phonetics, and time permitting, aspects of hearing and perception.

Topics of acoustic phonetics include, but are not limited to, understanding the physics of speech sounds, identifying acoustic signals corresponding to various sounds used in human speech, and applying phonetic theories in analyzing data that are previously unknown to learners.

Data for acoustic analyses often come from pre-recorded materials based on major languages such as English or Japanese, or from materials recorded by the learners themselves. Recording one's own voice, which can in turn be used for applying phonetic theories, is an effective and motivating way of teaching phonetics. Drawbacks of this method appear when teachers evaluate the learning outcome for the learners. Individual-specific speech materials recorded on non-uniform devices often lead to a

difficult situation when teachers want to be fair in the evaluation process because penalizing recording quality is not the goal of a phonetics course.

I propose that using field data based on understudied languages serves the purpose of applying phonetic knowledge to unknown data. Moreover, a uniform field data set also generates comparable responses to an assignment that can be used to fairly judge learners' achievement.

From the perspectives of learners, concepts of acoustic phonetics such as voice onset time (vot), duration, formants, pitch, etc. may be daunting at first. However, coupling these concepts with data from unknown languages that is not usually accessible to learners has resulted in raised interest and stronger motivation in my phonetic classes.

In this paper, I will discuss three representative case studies for phonetics studies; Dzongkha data is used for introducing voice onset time, Miyako Ryukyuan data is used for teaching characteristics of nasals, and Xitsonga data is used for demonstrating how to teach vowel quality plots.

2. Voice onset time in Dzongkha

2.1. Laryngeal contrast in Dzongkha

Dzongkha is a Tibeto-Burman language spoken by about 640,000 people in the Kingdom of Bhutan. Dzongkha data has been obtained as part of an ongoing project on three Tibeto-Burman languages: Dzongkha (Bhutan), Dränjongke (India), and Tamang (Nepal).

As in (1), plosives in Dzongkha have a four-way laryngeal contrast [1, 2]. This contrast in Dzongkha is unique because most languages lack the devoiced category in (1d).

- (1) a. voiceless unaspirated [pa]
 b. voiceless aspirated [pʰa]
 c. voiced [ba]
 d. devoiced [b̥a]

As an unknown language to most learners in my classes, Dzongkha illustrates how voice onset time (vot) is used to reliably distinguish the laryngeal

contrast. The advantage of using Dzongkha, instead of English, comes not only from the vot contrast between voiceless unaspirated and voiceless aspirated, but also from the strong presence of prevoicing in the voiced category [ba]. In English, the pronunciation of the letter ‘b’ is not accompanied by strong prevoicing.

In addition, the devoiced category is intriguing to students because they can hear the difference, but they cannot immediately identify what distinguishes the devoiced category from the voiceless unaspirated category. Depending on the level of knowledge of phonetics, some students may point out that the fundamental frequency (f_0) of the vowel following the devoiced plosive is lower than that of the vowel following the voiceless unaspirated. The devoiced category makes students aware of the inter-connectivity between phonetic categories; phonological contrast does not have a one-to-one correspondence to phonetic measurements.

2.2. Dzongkha data in the classroom

Learners already know how to use Praat and they know acoustic characteristics of plosives: closure followed by a burst. After introducing the basics of Dzongkha, learners are asked to annotate the voiceless unaspirated stops using Praat as in Fig 1. Then, learners annotate voiceless aspirates as in Fig 2 and observe the difference between unaspirated and aspirated plosives. The annotation of the voiced category shows strong prevoicing (Fig. 3). It is an intuitive way of being exposed to the voiced category.

The last category is then assigned for annotation by learners, without providing any further information about the sound (Fig 4). At this point, learners have to annotate and identify a sound based on what they have learned until this point. Most students are puzzled by the similarity between the devoiced and voiceless unaspirated category. The difference is explained using the fundamental frequency (f_0), which is lower after a devoiced onset compared to after a voiceless unaspirated onset (see Fig. 5).

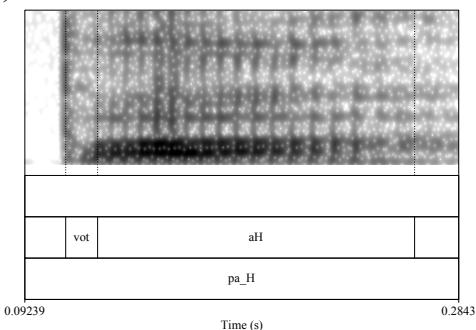


Figure 1: Voiceless unaspirated [pa] in Dzongkha.

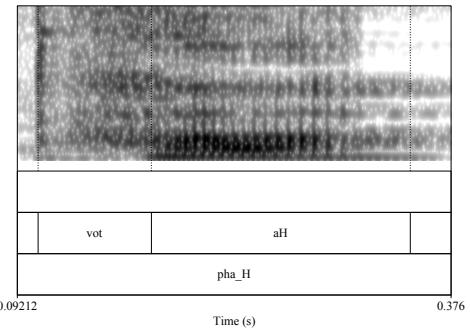


Figure 2: Voiceless aspirated [p^ha] in Dzongkha.

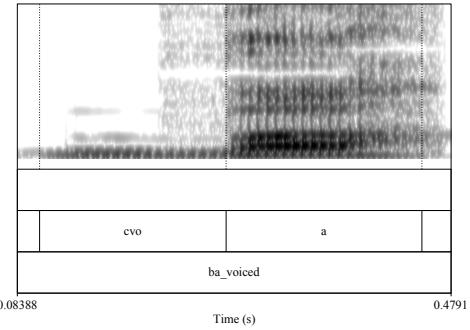


Figure 3: Voiced [ba] in Dzongkha; cvo means closure voicing.

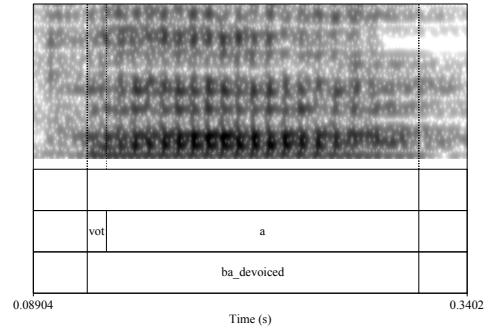


Figure 4: Devoiced [ba] in Dzongkha.

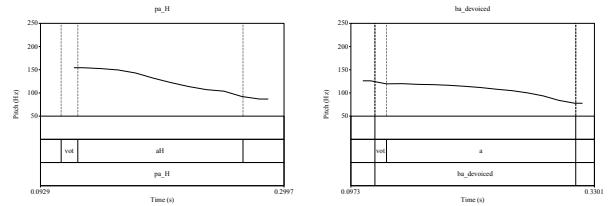


Figure 5: F0 of vowel following voiceless unaspirated [pa] (ca 150 Hz) and devoiced [ba] (ca 125 Hz) in Dzongkha.

Upon completing the Dzongkha assignment, learners will know how to identify voice onset time and prevoicing, and also know that identification of a sound requires considerations of multiple acoustic features. Learners’ knowledge will not be bound by the more common two-way laryngeal contrast found in English or Japanese.

3. Duration in Miyako Ryukyuan

3.1. Nasals in Miyako Ryukyuan

Miyakoan is a Ryukyuan language spoken in Miyako island in Okinawa. This language has a contrast between singleton nasals, geminate nasals and voiceless nasals [3]. This three-way contrast in nasals offers a showcase of understanding the relationship between acoustic duration and phonological length contrast. Moreover, Miyakoan offers a window into understanding sonorants that are not voiced (Table 1).

Table 1: Nasals in Miyakoan Ryukyuan: singleton nasals, geminate nasals and voiceless nasals

mami	'bean'	m:a	'mother'	m:adi	'let's step on'
nabi	'pan'	n:a	'conch'	n:a	'rope'
ŋkla:n	'old days'	ŋ:	'sweet potato'	ŋ:gamata	'will step on'

The contrast between singleton and geminate is often introduced with the plosive category in Japanese: singleton in [oto] 'sound' versus geminate in [otto] 'husband'. Learners readily understand this contrast by pointing out that geminates are sokuon written with a small 'tsu' symbol. When faced with acoustic features, they become puzzled, because the contrast between singleton and geminate is manifested in the closure duration. In other words, the duration of the absence of an acoustic signal is a determining factor for distinguishing geminates from singletons.

Introducing the contrast induced by the acoustic duration of nasality offers an intuitive way of understanding the singleton versus geminate contrast because learners are facing information that is present in the acoustic signal.

3.2. Miyako data in the classroom

Learners have been taught how to make boundaries and annotate intervals using Praat. They also know acoustic characteristics of nasals: dampening of the amplitude and antiformants. Three sound files are provided to learners for investigation. The target words are surrounded by a frame sentence, which is also provided to learners. In the task, learners need to annotate [mipana] 'face' for a singleton nasal, [mmidu:] 'being ripened' for a geminate nasal, and [nna] 'traditional rain coat' for a voiceless nasal (Figures 6, 7, 8 respectively).

The results of the measurements are 98 ms for a singleton nasal and 204 ms for a geminate nasal. In addition to seeing the difference in nasal duration, learners will also realize that geminates are not exactly twice as long as singletons.

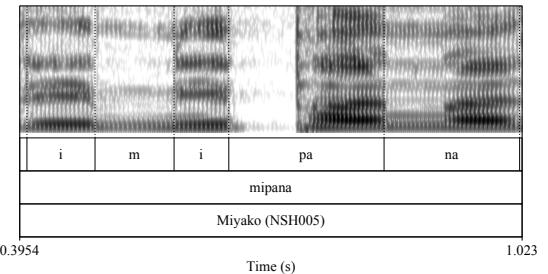


Figure 6: A singleton nasal in Miyako.

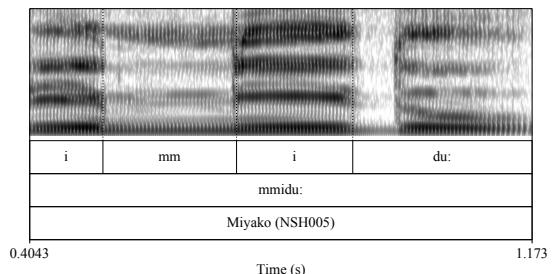


Figure 7: A geminate nasal in Miyako.

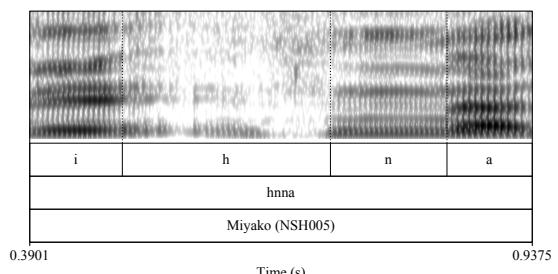


Figure 8: A voiceless nasal in Miyako.

Voiceless nasals in Miyako have a weak glottal fricative (similar to [h]) preceding the nasal part. Through voiceless nasals, learners can readily identify the nasal portion preceding a vowel. Thereafter, they learn that nasals are not always voiced, and a voiceless nasal is a sound with two acoustic features in sequence.

By completing the task on Miyako nasals, learners will know that the distinction between singletons and geminates can be measured using duration. This knowledge is the basis for understanding the geminate plosives in a language, which has a long duration as a portion that has no sound. Moreover, learners have access to a typologically uncommon sound (a voiceless nasal) that consists of both an [h] and a nasal together in a single sound.

4. Formants and Xitsonga

4.1. Vowels in Xitsonga

Xitsonga is a southern Bantu language spoken mainly in South Africa, Mozambique and Zimbabwe [4, 5, 6]. Xitsonga has a five-vowel system: [i, e, a, o, u]. Five words with an initial [n], and with a following vowel that varies, are selected for this activity (see (2)).

- (2) a. [nala] ‘enemy’
 b. [neŋe] ‘leg’
 c. [nija] ‘cricket’
 d. [nomo] ‘mouth’
 e. [nunju] ‘porcupine’

Xitsonga vowels are used to represent formant values and then to understand the relationship between formant values and vowels. Figure 9 represents how the relationship between the first formant (F1) and the second formant (F2) defines types of vowels.

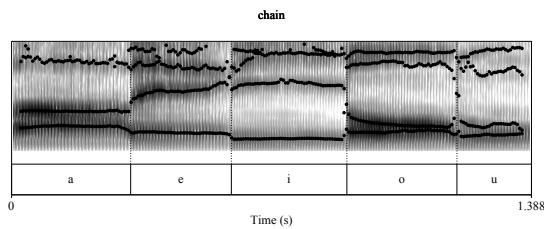


Figure 9: Vowel formants taken from five Xitsonga words: [nala] ‘enemy’, [neŋe] ‘leg’, [nija] ‘cricket’, [nomo] ‘mouth’ and [nunju] ‘porcupine’.

4.2. Xitsonga data in the classroom

Learners have been taught how to use Praat Pictures to draw various features in a single panel. In the Xitsonga task, learners will annotate vowels, extract a vowel from each word, concatenate the vowels, and then draw spectrograms and formants using Praat Pictures. The task was originally designed for learners of phonetics in a university classroom in Japan. Japanese has a five-vowel system, so working on Xitsonga gives access to a different language with a five-vowel system.

The task represents formants of the five vowels. The low vowel [a] has a larger F1 value and the high vowel [i] has a smaller F1 value. The distance between F1 and F2 provides acoustic cues regarding the front-back dimension. The front vowel [i] has a larger F2 – F1 distance, whereas the back vowel [o] has a smaller F2 – F1 distance.

Learners who successfully complete this task apply the knowledge of formant values to an unknown language (Xitsonga), and learn that the acoustic dimensions described for English vowels

and/or Japanese vowels holds for a Bantu language such as Xitsonga.

5. Conclusion

Acoustic phonetics is a rather challenging area in phonetics due to the requirements for the knowledge of basic physics as well as the need to use Praat software. If accessible, utilizing acoustic data from understudied languages complements a common format in a phonetics course where the bulk of the data come mainly from major languages such as English.

This paper has shown how data from Dzongkha, Miyako Ryukyuan and Xitsonga help learners to apply phonetic knowledge such as voice onset time, duration and formants to previously unknown languages. Teaching phonetics is not a simple task, but resources such as the ones introduced here could alleviate learners’ fears when they have to face acoustic signals for understanding phonological contrasts.

6. Acknowledgements

Dzongkha data was obtained from a grant by the Strategic Japanese-Swiss Science and Technology Programme of JSPS and SNSF. Miyako Ryukyuan data come from a Kashima Foundation grant to Yuuka Hayashi. Xitsonga data come from a Kakenhi(C) grant 16K02641. Thanks to Jeremy Perkins who proofread this paper. For those who are interested in sound files discussed in this paper, please contact me individually.

7. References

- [1] Lee, S. J., Kawahara, S. 2018. The phonetic structure of Dzongkha: a preliminary study *Journal of the Phonetic Society of Japan* 22(1), 13-20.
- [2] Perkins, J., Lee, S. J., Kawahara, S., Monou, T. 2018. Consonants and tones: A view from two Tibeto-Burman languages, *The 156th meeting of Linguistic Society of Japan*.
- [3] Hayashi, Y., Aso, R., Kawahara, S., Lee, S. J. 2017. Partially devoiced geminate nasals in the Ikema dialect of Miyakoan: A preliminary study on their durational properties, *Seoul International Conference on Speech Sciences (SICSS)*.
- [4] Lee, S. J. 2015. Cumulative effects in Xitsonga: high-tone spreading and depressor consonants, *Southern African Linguistics and Applied Language Studies* 33(3), 273-290.
- [5] Bennett, W., Lee, S. J. 2015. A surface constraint in Xitsonga: *LI, *Africana Linguistica* 21, 4-27.
- [6] Lee, S.-I., Kawahara, S., Lee, S. J. 2015. The Whistled fricatives in Xitsonga: Its Articulation and Acoustics, *Phonetica* 71(1), 50-81.